Journal of Biomolecular Techniques
13:43–48 © 2002 ABRF

COMMUNICATIONS

Protein Identification by Automated Nanospray Mass Spectrometry—"Zoomscan Walking"

*Matt J. Sweeney,[a] Ben C.B. Ko,[b] Stephen S.M. Chung,[c] and Christoph W. Turck[d]**

[a]Matt Sweeney Mass Spec Consulting, Los Altos, CA; [b]Department of Chemistry and Open Laboratory of Chemical Biology of the Institute of Molecular Technology for Drug Discovery and Synthesis, and [c]Institute of Molecular Biology, The University of Hong Kong, Hong Kong, China; [d]Howard Hughes Medical Institute, Department of Medicine and Cardiovascular Research Institute, University of California San Francisco, San Francisco, CA

Nanospray mass spectrometry is an important tool for high-sensitivity protein identification. A disadvantage of nanospray infusion is that the operator must search for peptide ions from a protein digest for subsequent collision-induced fragmentation. In order to make the nanospray method less operator-dependent, we have implemented an automated procedure using commonly available tools. Protein digests are subjected to unattended nanospray analysis using an automatic "zoomscan walking" procedure. The identification of eight proteins from an SDS gel using automated "zoomscan walking" is shown as an example of the utility of the method.

Key Words: nanospray, mass spectrometry, zoomscan.

ADDRESS CORRESPONDENCE AND REPRINT REQUESTS TO: Christoph W. Turck, Max Planck Institute of Psychiatry, Kraepelinstr. 2, D-80804 Munich, Germany (phone: 49-89-30622-317; fax: 49-89-30622-200; email: turck@mpipsykl.mpg.de).

*CURRENT ADDRESS: Max Planck Institute of Psychiatry, Molecular, Cellular, Clinical Proteomics, Munich, Germany.

MATERIALS AND METHODS

Cell Culture and Media

HeLa cells were grown in Dulbecco's Modified Eagle's medium supplemented with 10% fetal bovine serum, according to supplier instruction (American Type Culture Collection, Rockville, MD).

Preparation of Affinity Column

A cDNA fragment encoding the first 380 amino acid residues of the osmotic response element binding protein KIAA0827[1] was cloned in-frame into the glutathione-S-transferase (GST) fusion vector pGEX-5T (Amersham Biosciences, Piscataway, NJ). The recombinant protein was expressed in the bacterial host BL21 (Amersham Biosciences) by IPTG induction for 4 h. The recombinant GST fusion protein, GST-$P_{1\text{-}380}$, was purified to apparent homogeneity using a glutathione Sepharose column (Amersham Biosciences) and coupled to CNBr-activated Sepharose according to supplier instruction (Amersham Biosciences).

Preparation of Cell Extracts and Protein Purification

HeLa cell cytoplasmic extracts were prepared by scraping confluent cultures into ice-cold lysis buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 0.5% Triton X-100) supplemented with 0.5 mM DTT and protease

inhibitor mixture (Roche Diagnostics, Indianapolis, IN). A total of 700 mg cytoplasmic extracts were preincubated with Sepharose for 1.5 h at 4°C. The extracts were filtered and applied to the GST-$P_{1\text{-}380}$ Sepharose affinity column at a flow rate of 1 mL/min. After extensive washing with phosphate-buffered saline (PBS) the column was eluted with a linear NaCl gradient (0–1 M) and the eluent collected in 1-min fractions. Proteins in the individual fractions were precipitated with 30% trichloroacetic acid (TCA), the pellets washed with acetone and air-dried. The precipitated proteins were then subjected to SDS-PAGE in a 10% gel and stained with Coomassie brilliant blue. After destaining, the gel bands of interest were cut out with a razor blade and the proteins subjected to an in-gel digest with 1 µg of endoproteinase Lys-C (Roche Diagnostics).[2] Peptide mixtures derived from the individual digests were desalted and eluted in a 1.5-µL volume into a nanospray capillary (MDS Proteomics Protana, Odense, Denmark) as described.[3]

Protein Identification

All experiments were carried out with an LCQ Classic iontrap mass spectrometer (ThermoFinnigan, San Jose, CA) equipped with a nanospray source (MDS Proteomics Protana). For protein identification, spectral MS/MS data were converted to "dta files" and searched against the nonredundant database using SEQUEST (John Yates, III and Jimmy Eng, University of Washington, Seattle, WA).[4] The "Mass Sequencing Gel" display was generated with the help of the Xcalibur 1.2 software (Thermo Finnigan).

RESULTS

HeLa cytoplasmic cellular extracts underwent a purification procedure that is outlined in Materials and Methods. An affinity resin with the osmotic response element binding protein KIAA0827 (Genbank Accession AB020634)[1] was employed, and in the final purification step bound proteins were eluted from the affinity column with a linear NaCl gradient. The proteins from each chromatography fraction were then concentrated by TCA precipitation and subjected to SDS-PAGE (Fig. 1).

After staining the gel, individual protein bands were cut out and subjected to in-gel digestion with endoproteinase Lys-C. The resultant peptide mixtures were analyzed by nanospray mass spectrometry[3] using automated "zoomscan walking" data-dependent scans (Fig. 2). The results of the analysis are presented in Tables 1 and 2.

The experiment method allows a large mass range to be automatically scanned and zoomscan and MS/MS data obtained automatically. The original automated "zoomscan walking" method utilized the parent ion mapping feature of the "gizmos" developer's package in the tune view of the Xcalibur software. The method forces the mass spectrometer to scan from low to high mass in units of the defined parent scan window (in our case 7 amu). Subsequently, a zoomscan is automatically triggered above a very low threshold that was specified (4000 counts). The zoomscan window was 10 amu. Two or three MS/MS scans are finally triggered on the most intense ions of the respective scans. A single microliter of protein digest sample typically generated 400–500 MS/MS scans producing a 3–4-Mb data file. The tune file was saved with the parent ion mapping feature of the "gizmos" enabled. This feature, in combination with the data-dependent triple play function, results in the "zoomscan walking" method. The instrument starts at the low-mass parent ion mapping setting and proceeds up to high mass in 7-amu increments. In each window, the two or three largest ions are used for MS/MS (Fig. 2). Sample volume, amount of protein, and the mass range that is desired are all factors that influence the selection of the parameters.

A map display of the entire data file of a "zoomscan walking" experiment is shown in Figure 3. The *x* axis displays the time and the *y* axis the mass/charge *(m/z)* values. The gray scale represents the signal intensities. The experiment begins on the left and continues to the right. At time zero "zoomscan walking" occurs in the 500-amu region and identifies candidate ions for MS/MS. These MS/MS scans cover a wide mass range while the zoom- and full-scan data focus on a narrow range (10 amu and 7 amu, respectively), resulting in the apparent "curve" across the picture. The slope of the curve is predominantly related to the speed required to fill the iontrap for the MS/MS stages. The slope is steeper when there are plenty of ions around, which requires less time to fill the iontrap. The width of the "bands" is also indicative of the length of time required to fill the iontrap. The more intense MS/MS spectra tend to appear as thinner bands and the less intense ones as thicker.

The choice of charge state is also apparent. All multiply charged ions have data recorded above and below the parent ion *m/z* value, which lies on the curve. The "1/3 low-mass cut-off rule" is visible in the data set as the lower limit of the MS/MS spectra, which tends to slowly increase with the parent mass. Based

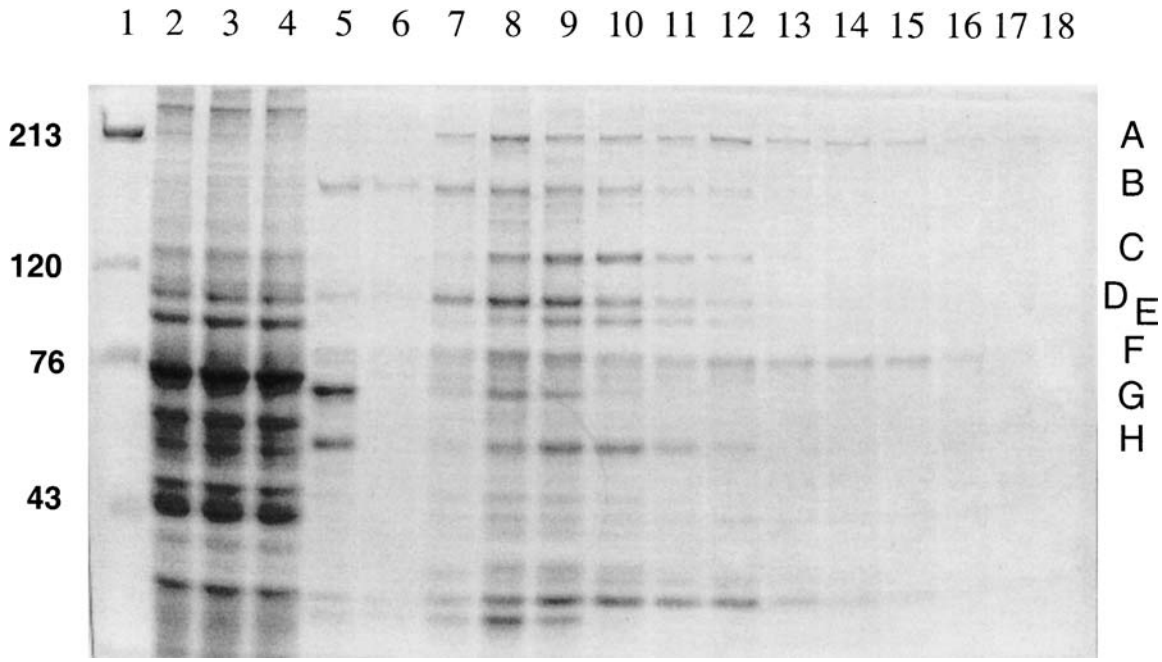


FIG. 1

SDS gel electrophoresis of affinity chromatography fractions. Proteins were precipitated with TCA, washed with acetone, and dissolved in SDS sample buffer. After electrophoresis the gel was stained with Coomassie brilliant blue. Gel bands containing proteins of interest that were cut out are indicated with letters on the right. The molecular weights of marker proteins are indicated on the left of the gel. *Lane 1:* molecular weight markers; *lanes 2–4:* crude ion exchange fraction that was loaded onto the affinity column; *lanes 5–18:* affinity chromatography fractions.

on its resemblance with classical DNA sequencing gels we have termed this map "mass sequencing gel."

We also implemented a "dynamic exclusion zoomscan walking" procedure. This method uses the standard "triple play" mode modified by the dynamic exclusion feature of the Xcalibur software. "Dynamic exclusion" builds a table of masses for which MS/MS spectra have been obtained. The process involves scanning the full mass range; finding the largest intensity ion which has not been included on the list, and performing a zoomscan and MS/MS; adding it to the exclusion mass list; and then repeating the cycle. The masses of common contaminants, e.g., enzyme autolysis fragments and plasticizers, can be added to a static table of exclusion masses. This method is fairly robust with respect to the commonly observed variations in relative ion intensities seen in spray ionization methods. The repetitive cycling of the dependent scans combined with dynamic exclusion means that, eventually, any ion of reasonable intensity will have MS/MS spectra recorded on it. The functionality and coverage of the "gizmos"-style method can be replicated by running a set of analyses, each covering a limited mass range to avoid the constraint imposed by the 25-element limit in the exclusion mass list. In both cases we utilized high numbers of microscans (9–16) and long maximum inject times (1000 ms) for the MS/MS and zoomscans. This can result in long scan times but has little impact on our nonchromatographic system.

TABLE 1

Proteins Identified in Gel Bands from Figure 1 by Automated "Zoomscan Walking" Method

Gel band	Protein
A	Talin
B	Clathrin heavy chain
C	Heterogeneous nuclear ribonucleoprotein U
D	Ewing sarcoma breakpoint region 1 RNA-binding protein
E	Keratin type II cytoskeletal
F	Fus-like protein
G	Probable RNA-dependent helicase P68
H	HSPC117

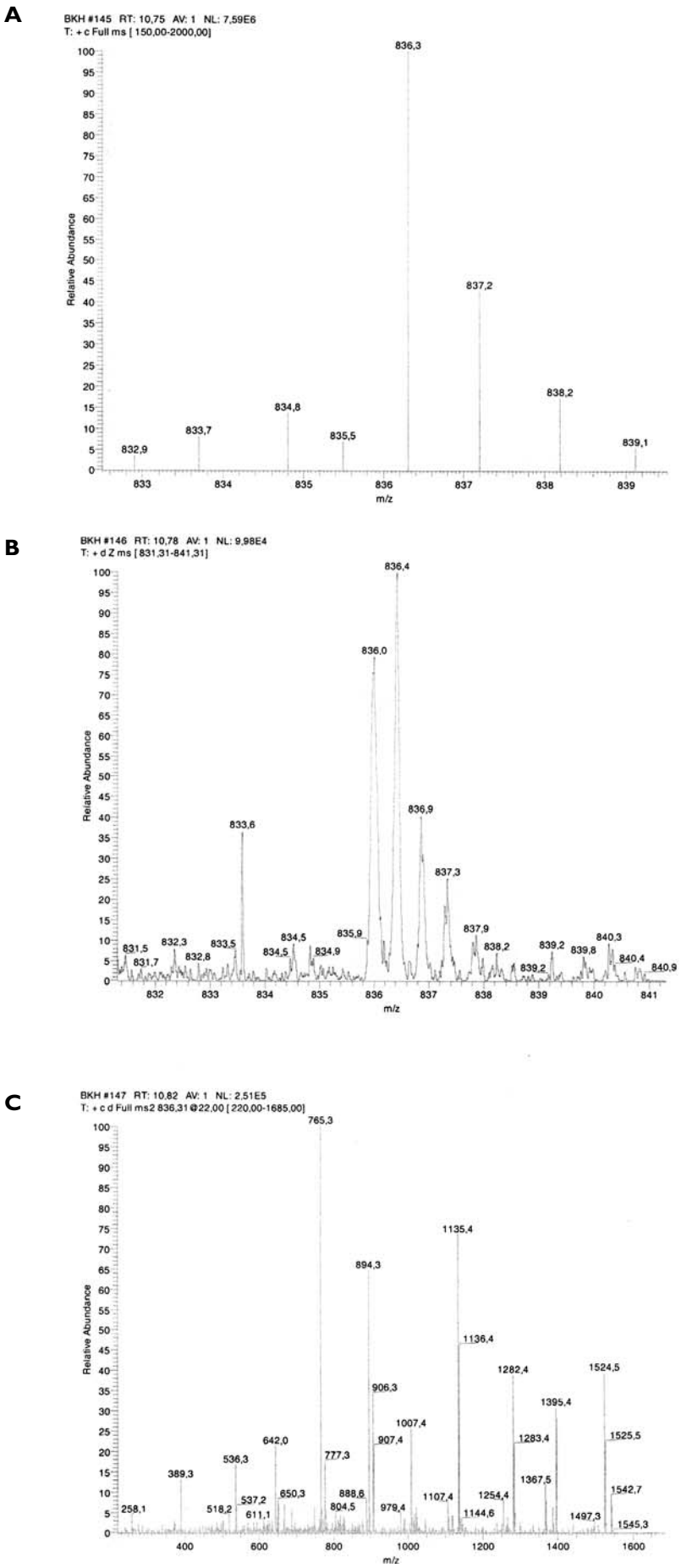
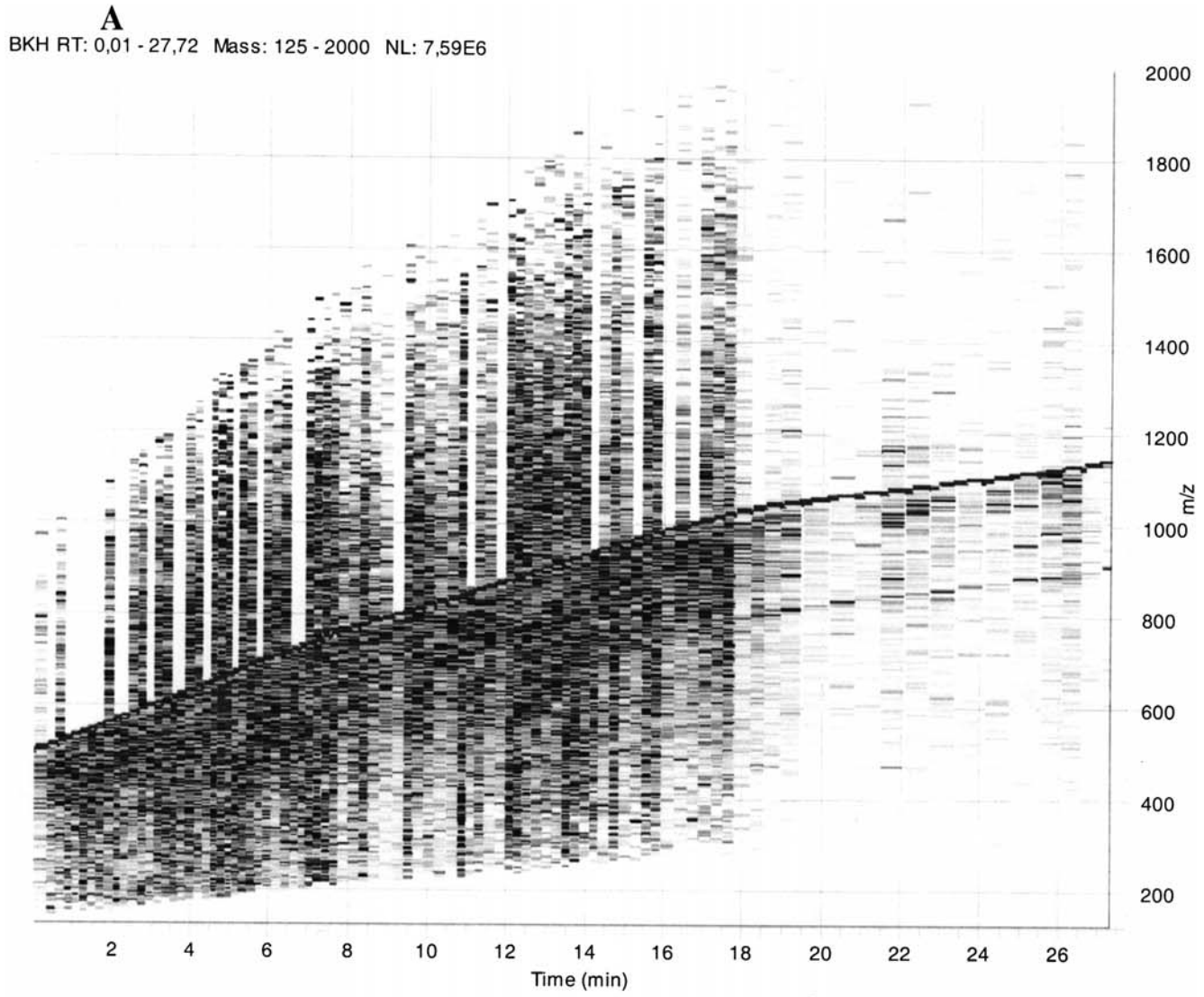


FIG. 2

Example of "zoomscan walking" method. **A:** Full-scan with a 7-amu window, followed by **B:** data-dependent zoomscan with a 10-amu window, and **C:** data-dependent MS/MS scan. The most intense ion of the full scan is subjected to a zoomscan followed by an MS/MS scan.

FIG. 3

"Mass sequencing gel"—a map display of a "zoomscan walking" experiment. The *x* axis represents the time of the run, the *y* axis the mass over charge values. The gray scale reflects the intensities of the signals; the width of the bands is an indication of the length of time required to fill the iontrap. **A:** Full display from 0 to 27.7 min. **B:** Zoom display from 8.5 to 12.6 min.


A
BKH RT: 0,01 - 27,72 Mass: 125 - 2000 NL: 7,59E6
2000
1800
1600
1400
1200
1000
800
600
400
200
m/z
2
4
6
8
10
12
14
16
18
20
22
24
26
Time (min)

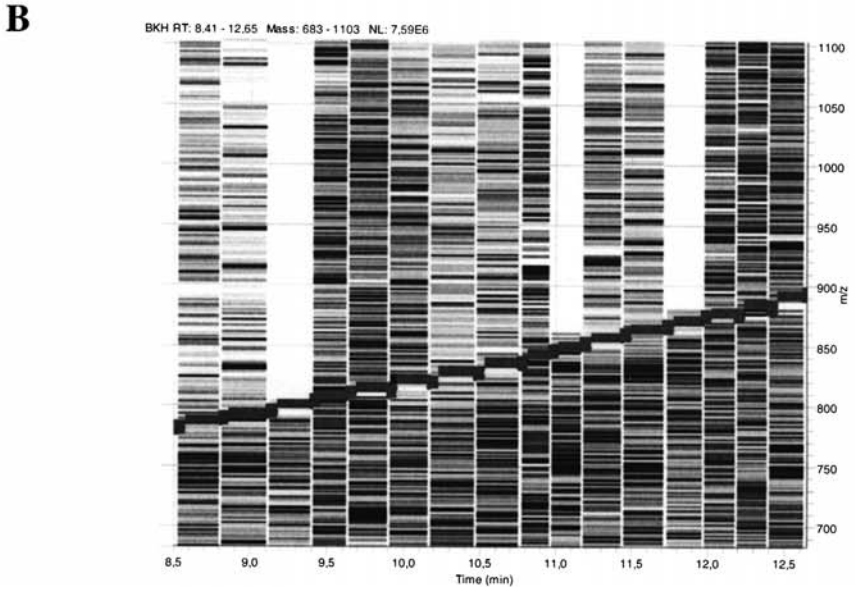

B
BKH RT: 8,41 - 12,65 Mass: 683 - 1103 NL: 7,59E6
1100
1050
1000
950
900
850
800
750
700
m/z
8,5
9,0
9,5
10,0
10,5
11,0
11,5
12,0
12,5
Time (min)

TABLE 2

SEQUEST Search Results of Automated "Zoomscan Walking" Experiment of Band A (Talin) Digest

Sequence position	Amino acid sequence
2351–2361	SIAAATSALVK
325–334	LVPRLLGITK
1321–1332	ALSTDPASPNLK
2532–2541	FLPSELRDEH
2044–2063	LAQAAQSSVATITRLADVVK
2362–2375	AASAAQRELVAQGK
1767–1780	TLAESALQLLYTAK
1026–1040	NLGTALAELRTAAQK
1416–1431	NGNLPEFGDAIATASK
2477–2491	AAAFEDQENETVVVK
923–943	QAAASATQTIAAAQHAASAPK
2044–2063	LAQAAQSSVATITRLADVVK
593–613	LLAALLEDEGGNGRPLLQAAK
2064–2085	LGAASLGAEDPETQVVLINAVK
307–316	TYGVSFFLVK
1321–1332	ALSTDPASPNLK
2362–2375	AASAAQRELVAQGK

Protein Coverage for Talin:
211/2541 = 8.3% by amino acid count
21608/269831 = 8.0% by mass

CONCLUSIONS

Automated "zoomscan walking" is an easy and time-saving way to reduce the tedious manual work of MS/MS on nanospray samples. It avoids the requirement of an HPLC and can identify proteins below 50 fmol/μL in an automatic analysis mode. The method may be implemented in a variety of ways, using the "parent ion mapping" or the "dynamic exclusion" features of the Xcalibur software. We prefer the "parent ion mapping" feature since ions of low intensity experience less discrimination and hence are better represented in the final data set.

ACKNOWLEDGMENTS

This work was supported by the Howard Hughes Medical Institute and Areas of Excellence Scheme established under the University Grants Committee of the Hong Kong Special Administrative Region, China (Project No. AoE/P-10/01), The University of Hong Kong (Generic Drugs Research Program).

REFERENCES

1. Ko BCB, Turck CW, Lee KWY, Yang Y, Chung SSM. Purification, identification, and characterization of an osmotic response element binding protein. *Biochem Biophys Res Comm* 2000;270:52–61.
2. Hellman U, Wernstedt C, Gonez J, Heldin CH. Improvement of an "In-Gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. *Anal Biochem* 1995;224:451–455.
3. Wilm M, Shevchenko A, Houthaeve T, Breit S, Schweigerer L, Fotsis T, Mann M. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 1996;379:466–469.
4. Eng JK, McCormack AL, Yates JR 3rd. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 1994;5:976–989.